Second opinion needed: communicating uncertainty in medical machine learning

Abstract

There is great excitement that medical artificial intelligence (AI) based on machine learning (ML) can be used to improve decision making at the patient level in a variety of healthcare settings. However, the quantification and communication of uncertainty for individual predictions is often neglected even though uncertainty estimates could lead to more principled decision-making and enable machine learning models to automatically or semi-automatically abstain on samples for which there is high uncertainty. In this article, we provide an overview of different approaches to uncertainty quantification and abstention for machine learning and highlight how these techniques could improve the safety and reliability of current ML systems being used in healthcare settings. Effective quantification and communication of uncertainty could help to engender trust with healthcare workers, while providing safeguards against known failure modes of current machine learning approaches. As machine learning becomes further integrated into healthcare environments, the ability to say "I'm not sure" or "I don't know" when uncertain is a necessary capability to enable safe clinical deployment.